PCT/DK 2004/000195

Rec'd PCT/PTO 20 SEP 2005

10/549619

REC'D 21 APR 2004

WIPO PCT

Kongeriget Danmark

BEST AVAILABLE COPY

Patent application No.:	PA 2003 00430
Date of filing:	20 March 2003
Applicant: (Name and address)	Nuevolution A/S Rønnegade 8, 5 DK-2100 København Ø Denmark

Title: Ligational encoding

IPC: C 07 H 19/00

This is to certify that the attached documents are exact copies of the above mentioned patent application as originally filed.

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Patent- og Varemærkestyrelsen • Kongeriget Danmark

Patent- og Varemærkestyrelsen
Økonomi- og Erhvervsministeriet

14 April 2004

Bo Z. Tidemann

PATENT- OG VAREMÆRKESTYRELSEN

Title

LIGATIONAL ENCODING

Technical Field of the Invention

The present invention relates to a method for synthesizing a bifunctional complex comprising an encoded molecule and a template coding for chemical entities which have participated in the synthesis of the encoded molecule. The invention also relates to a library of different complexes, said library being obtainable by processing a plurality of different templates. The library of the invention may be suitable for identifying drugs.

Background of the Invention

Libraries of complexes comprising an encoded molecule as well as the template which has coded for the synthesis thereof are useful in finding new chemical compounds which may be used for therapeutic purposes, because the potential drug is connected to an identifier molecule or template, which may be decoded for identification of each chemical entity that has participated in the synthetic history.

Some attempts to form the complex comprising an encoded molecule as well as the template that codes for the chemical entity that has participated in the formation of the encoded molecule, were based on the split-and-mix principle known from combinatorial chemistry, see e.g. WO 93/06121 A1, EP 643 778 B1, and WO 00/23458. If several selection rounds are desirable or necessary the split-and-mix principle has the inherent disadvantage of requiring decoding between each selection round. The decoding step may be laborious and cumbersome because the templates usually are incorporated into a vector and then subsequently into a suitable host micro organism.

Other attempts have focussed on the formation of encoded proteins using the natural machinery of a cell and connecting the formed protein with the template nucleic acid that has coded for the amino acid components of the protein. Examples of suitable systems are phage display, *E. coli* display, ribosome display (WO 93/03172), and protein-mRNA-fusions (WO 98/31700). The genetic information of the nucleic acid, usually mRNA or DNA, may not necessarily be decoded between each round of selection to establish the identity of the chemical entities that has formed the protein because the

nucleic acid can be amplified by known means, such as PCR, and processed for the formation of a new library enriched in respect of suitable binding proteins.

Recently, new a method for encoding molecules has been suggested, which can be performed in several selection rounds without intermediate decoding, wherein the encoded molecule is not restricted to peptides and proteins. WO 02/00419 and WO 02/103008 disclose methods for preparing virtually any molecule connected to a template coding for chemical entities which have reacted to form the molecule. In short, a template segregated into a plurality of codons and a plurality of building blocks comprising a transferable chemical entity and an anticodon are initially provided. Under hybridisation conditions, the template and building blocks are annealed together and the chemical entities are subsequently reacted to form the molecule. The methods of the prior art are, however, restricted to reactions of the chemical entities which can be performed under hybridisation conditions. Hybridisation conditions generally imply aqueous solvents, moderate pH, and ambient temperature.

In one aspect of the present invention, it is the object to provide an encoding method, which expands the possible chemical reactions available for producing encoded molecules.

Summary of the Invention

The present invention discloses a method for synthesising a bifunctional complex comprising an encoded molecule and a template coding for one or more chemical entities which have participated in the synthesis of the encoded molecule, the method comprising the steps of

i) providing a) a template comprising one or more codons, b) one or more building blocks having an anticodon associated with a chemical entity, and c) a nucleic acid sequence associated with a reactive site,

ii) contacting the tempate with the one or more building blocks under conditions allowing for hybridisation between codons and anticodons,

iii) ligating at least one anticodon of a building block to the nucleic acid sequence associated with the reactive site, and

iv) reacting the chemical entity of the ligated building block with the reactive site under conditions where the ligation product is single stranded, to obtain a template-encoded reaction product.

The media for performing a reaction product is of crucial importance for the progress of the chemical reaction. As an example, many chemical reactions cannot be effectively conducted in aqueous solvents because the reactants are not sufficient soluble. Moreover, when nucleic acids are present, a lipophilic reactant may prefer, in an aqueous solvent, to be located in or in the vicinity of the double helix and therefore not accessible for another reactant dissolvable in the solvent. The present invention provide a solution to this problem by allowing a covalent link between the chemical entity and the reactive site to be reacted with the chemical entity, thereby allowing non-hybridising conditions to be present during the reaction. The upper limit for the conditions applied during the reaction may be the degradation of the nucleic acids. However, nucleic acids are stable molecules withstanding high temperatures, extreme pH, most organic solvents etc.

Numerous chemical reactions are compatible with DNA chemistry. However, only a limited number of such reactions are compatible with the presence of a DNA duplex formed between the anticodon of a building block and the template. In an aspect of this invention, the separation of conditions for performing the genetic information exchange step and the chemical reaction step ensure that many additional chemical reactions which are not compatible with a DNA duplex is accessible to the experimenter. In addition, this technology has the potential of increasing the speed, specificity and cost-efficiency of template programmed chemical reactions.

As acknowledged by those skilled in the art a plethora of means exist for the denaturation of DNA duplexes or the removal of a single strand of a duplex such as heat, alkali or acid, denaturant such as urea, formamide, GdHCl, ethanol, isopropanol, methanol, hygroscopic and/or organic solvents or any combinations of the above as well as nucleases or molecular handles enabling the specific removal or partial removal of a template strand(s).

Generally, the template comprises one or more codons. A single codon may be sufficient when the template or a nucleic acid hybridised to the template comprises a reactive site. Usually, the template comprises more than one codon to allow for a sufficient diverse encoded molecule. In a preferred aspect of the invention the template comprises 2-100 codons. Templates comprising more than 100 codons may be used but is

generally not necessary to afford the desired diversity of a library of complexes. In a preferred aspect of the invention the template comprises 3-20 codons.

The codon is a recognition unit that can be recognized by an anticodon. A variety of different kinds of recognition units exist in nature. Examples are antibodies which are recognized by an epitope, proteins which are recognized by another protein, mRNA which recognizes a protein, and oligonucleotides which recognize complementing oligonucleotide sequences. In certain aspects of the present invention a codon is a sequence of nucleotides. Generally the codon has the ability to interact with an anticodon in a specific manner, which allow for a specific recognition between a particular codon and anticodon pair. The specific pairing makes it possible to decode the template in order to establish the synthetic history of an encoded molecule. When the template comprises more than one codon, each member of a pool of building blocks can be identified uniquely and the order of the codons is informative of the synthesis step each member has been incorporated in.

The sequence of the nucleotides in each codon may have any suitable length. Generally it is preferred that each codon comprises two or more nucleotides. In certain aspects of invention each coding comprises 3 to 30 nucleotides, preferably 5 to 10 nucleotides.

The template will in general have at least two codons which are arranged in sequence, i.e. next to each other. Two neighbouring codons may be separated by a framing sequence. Depending on the encoded molecule formed, the template may comprise further codons, such as 3, 4, 5, or more codons, as indicated above. Each of the further codons may be separated by a suitable framing sequence. Preferably, all or at least a majority of the codons of the template are arranged in sequence and each of the codons is separated from a neighbouring codon by a framing sequence. The framing sequence may have any appropriate number of nucleotides, e.g. 1 to 20. Alternatively, codons on the template may be designed with overlapping sequences.

The framing sequence may serve various purposes. In one setup of the invention, the framing sequence identifies the position of a codon. Usually, the framing sequence either upstream or downstream of a codon comprises information which allows determination of the position of the codon. In another setup of the invention, the frames have

alternating sequences, allowing for additions of building blocks from two pools in the formation of a library.

The framing sequence may also or in addition provide for a region of high affinity. The high affinity region may ensure that the hybridisation of the template with the anti-codon will occur in frame. Moreover, the framing sequence may adjust the annealing temperature to a desired level.

A framing sequence with high affinity can be provided by incorporation of one or more nucleobases forming three hydrogen bonds to a cognate nucleobase. An example of a nucleobase having this property is guanine and cytosine. Alternatively, or in addition, the spacer sequence may be subjected to back bone modification. Several back bone modifications provides for higher affinity, such as 2'-O-methyl substitution of the ribose moiety, peptide nucleic acids (PNA), and 2'-4' O-methylene cyclisation of the ribose moiety, also referred to as LNA (Locked Nucleic Acid).

The template may further comprise a priming region for initiating the ligation process. The priming region allows for a ligation primer to hybridize to the template using appropriate conditions. Suitably, the ligation primer is a nucleic acid sequence which may or may not be associated with a reactive site. In addition to one or more codons the template may comprise a flanking region. The flanking region can encompasses a signal group, such a flourophor or a radio active group, to allow a direct detection of the presence of the complex or a label that may be detected, such as biotin. When the template comprises a biotin moiety, a hybridisation event can be observed by adding stained streptavidine, such as streptavidine-phycoerythrin conjugate. In a particular embodiment, the flanking regions are present on each side of the coding sequences providing for an amplification reaction, such as PCR.

In a certain aspect of the invention the flanking region is complementary to the priming region allowing for a hairpin loop to be formed when suitable hybridisation conditions is present. Suitably, no coding regions are present between the flanking region and the priming region. The use of a hairpin loop allows for covalent attachment of the nascent encoded molecule to the template that has encoded the synthesis of said molecule. In a certain aspect of the present invention, the duplex formed by the flanking region and the priming region is recognized by a restriction enzyme as substrate. The cleavage of the double helix allows for the separation of the ligation product and the template.

The one or more building blocks used in accordance with the present invention comprise an anticodon and a chemical entity. The anticodon and the chemical entity can be associated by a direct or indirect covalent or non-covalent interaction. Suitably, the anticodon is covalently connected to the chemical entity, optionally through a suitable linker.

The chemical entities can generally be divided into three groups. A first group of chemical entities comprises scaffold molecules. A scaffold molecule may be a single reactive group or a chemical core structure, like a steroid, to be modified. Generally the scaffold remains attached to the anticodon throughout the formation of the encoded molecule, thereby forming an anchorage point for the encoded molecule. The scaffold molecule may comprise more than a single reactive group. Usually, the one or more reactive groups of the scaffold are recipient reactive groups, i.e. reactive groups capable of forming a chemical connection to another chemical entity.

A second group of chemical entities comprises chemical entities which are capable of being transferred to a recipient reactive group, e.g. a recipient reactive group of a scaffold. The chemical entity can be selectively cleaved from the remainder of the building block by a suitable process following the formation of a connection between the chemical entity and the recipient reactive group. The selective cleavage may be suitable because the formation of the chemical bond between the chemical entity and the recipient reactive group and the cleavage of the chemical entity from the remainder of the building block can proceed in two separate steps using optimal conditions for each step. Alternatively, the reaction proceeds in a single step, i.e. the chemical entity is simultaneously reacted with the reactive site and cleaved from the remainder of the building block. The latter method involving simultaneous reaction and cleavage may be preferred when a fast method for formation of a single encoded molecule or a library of molecules are envisaged.

According to the third group of chemical entities, one part of an affinity pair is applied. Suitable examples of the one part of the affinity pair is biotin and dinitrophenol. Biotin can be selectively recognized by avidine or streptavidine and dinitrophenol can be selectively recognized by an antibody raised against that epitope. The incorporation of one part of an affinity pair into the ligation product may be useful in an immobilisation

process. As an example, the bifunctional complex comprising the encoded molecule, may be recovered following the partition step simply by adding the immobilized second part of the affinity pair.

The reactive groups appearing on the anticodons may in some embodiments be protected at the 3' or 5' end, because it may be desirable to be able to direct the incorporation of the individual building blocks. It may also be desirable to have the building block immobilized before the incorporation in the ligation product. The immobilisation may be achieved attaching the protection group of the anticodon to a solid support, such that the protection group appears between the anticodon and the solid support. The advantages of immobilising the nascent building block to a solid support is that it is possible to produce the final building block while remaining connected to the solid support. As is well known to the skilled organic chemist, solid support synthesis affords many advantageous over liquid reactions. The protection group is in an aspect of the invention photocleavable and preferably cleavable by exposure to UV light. In a preferred embodiment, a phosphate group is formed at the 5' end of the anticodon by deprotection, converting the anticodon to a substrate of a ligase. When a ligation primer or a nascent ligation product exposes a 5'-phosphate and the anticodon of a building block to be incorporated is able to hybridise next to the nucleotide comprising the 5'-phosphate, the 3'-end of the anticodon can be ligated to the 5'-end of the primer or nascent ligation product by a suitable ligase. Subsequently, the ligation product is exposed to a condition which deprotects the 5'-end of the anticodon, thereby providing a 5'-phosphate group of the nascent ligation product, which may be used in a subsequent incorporation of building blocks.

The nucleic acid sequence associated with the reactive site may involve a covalent or non-covalent, such as hybridisation, attachment between the sequence and the reactive site. Suitably the reactive site is covalently attached to the template. The reactive site may be part of a scaffold molecule or may be chemical entity according to the second group described above. In one aspect of the invention the nucleic acid sequence associated with a reactive site is a building block. The formation of an encoded molecule according to the present invention generally implies that a first anticodon of a building block is ligated to a primer complementing a priming sequence of the template. In some embodiments of the invention the primer may be absent and the ligation product is formed by ligating building blocks together. However, when enzymatic ligation is

encountered a ligation primer is generally used, even though it is possible simply to ligate building blocks together with the application of a ligation primer. In one aspect of the invention, the primer is covalently connected to the template, thereby forming a covalent connection between the anticodon and the template. The covalent connection may be formed by chemically cross-linking the strands or by connecting the primer trough a hairpin loop to the template.

The anticodon of a building block may be part of an oligonucleotide further comprising a sequence complementing a framing sequence of a template or a part thereof. The complementing framing sequences makes it possible for anticodons to recognise specific positions of codons on the template. As explained above the framing sequences and thus the complementing sequences may be alternating to allow for two different pools of building blocks to be added. The incorporation of building blocks can occur stepwise or two or more building block may be incorporated in a ligation product in the same ligation step. Stepwise ligation of building blocks may be desirable when the encoded molecule is formed by stepwise reacting the chemical entities arriving with newly incorporated building blocks. Incorporation of two or more building blocks may be useful when orthogonal chemical strategies are used, i.e. a reactive group of a chemical entity can react with one reactive group of a scaffold only, whereas other chemical entities comprises reactive groups which may react with distinct reactive groups of a scaffold.

The nucleic acid sequence associated with a reactive site used in the present invention may be comprised of a nascent encoded molecule associated with a ligation product. Subsequently, in one setup of the invention, an anticodon of a building block is ligated to a preceding incorporated anticodon and the chemical entity is reacted. According to another embodiment two or more building blocks are hybridised to the template and subsequently ligated together to form a ligation product.

Generally, a building block is hybridised next to another building block or a primer in order for a ligation to proceed. However, in some aspects of the invention, it may be suitable to have a building block hybridised in a position spaced one or more nucleotides from another building block, nascent ligation product or primer. A spacer nucleotide can be used for joining the building block with the preceding building block, ligation product, or the primer.

In an aspect of the invention, a building block being immobilized on a solid support is hybridised to a codon and subjected to a ligation reaction, followed by a detachment of the building block from the solid support, as explained elsewhere herein.

Different approaches for ligating anticodons, primers and/or nascent ligation products can be applied. According to a first approach the anticodon is ligated to a nucleic acid by chemical means. The chemical means may be selected from various chemistries known to the skilled man. Examples of chemical ligation methods include:

a) a nucleic acid, such as a first anticodon, comprising a 3'-OH group and a second nucleic acid, such as a second anticodon comprising a 5'phosphor-2-methylimidazole group. The 3'- and 5' reactive group are reacted to form a phosphodiester internucleoside linkage,

b) a nucleic acid, such as a first anticodon, comprising a phosphoimidazolide group at the 3'-end and a second nucleic acid comprising a phosphoimidazolide at the 5'-end, which are reacted to form a phosphodisester internucleoside linkage,

c) a nucleic acid, such as a first anticodon comprising a 3'-phosphorothioate group and a second nucleic acid sequence comprising a 5'-iodine, which are reacted to form the internucleoside linkage 3'-O-P(=O)(OH)-S-5', and

d) a nucleic acid, such as a first anticodon comprising a 3'-phosphorothioate group and a second nucleic acid, such as a second anticodon comprising a 5'-tosylate, which are reacted to form the internucleoside linkage 3'-O-P(=O)(OH)-S-5'.

In a preferred aspect of the invention, an enzyme is used for ligating an anticodon to a nucleic acid. The enzymes capable of ligating two nucleic acids together are generally referred to as ligases. Preferred ligases are selected from the group consisting of DNA ligase, and RNA ligase. The DNA ligase may be selected among the group consisting of Taq DNA ligase, T4 DNA ligase, T7 DNA ligase, and E. coli DNA ligase. In some aspects of the invention enzymatic ligation is preferred because a higher specificity generally is obtained and shorter anticodons may be used.

Following the ligation step the reaction of chemical entities is conducted at conditions where the ligation product is single stranded. A single stranded ligation product may be obtained in various ways. In one aspect of the invention, the single stranded ligation product is obtained using denaturing conditions. The denaturing conditions may i.a. be

obtained by using a media selected from organic solvents, aprotic solvents, acidic solvents, denaturants, and alkaline solvents. In another aspect of the invention the denaturing conditions are obtained by heating to a temperature above the melting temperature of the duplex. The single stranded ligation product may also be obtained by degrading the template.

The template can be degraded by various means, e.g. by providing an DNA template and an RNA ligation product and treating the DNA:RNA duplex with RNAseH, RNAseA, RNAse 1, weak alkaline conditions (pH 9-10), or aqueous Pb(Ac)2; by providing a DNA template comprising a thiophosphate in the internucleoside linker and an DNA or RNA anti-codon ligation product, and subsequent treating with aqueous iodine; or providing a DNA or RNA ligation product and a DNA template comprising an uracil nucleobase, treating with uracil-glycosylase and subsequent weak acid.

In another aspect the single stranded ligation product is obtained by removing the template. The template may be removed by a process comprising cleaving a bond between the ligation product and the template, subjecting to denaturing conditions and separating of the template. The bond between the ligation product and the template may be cleaved by a restriction endonuclease.

In certain aspects of the invention, the template is separated from the ligation product by a process which involves providing the template or the ligation product with a first part of an affinity pair. The first part of the affinity pair may be biotin or a similar moiety. The template or the ligation product having appended a biotin moiety can be bound to avidine or streptavidin immobilized on a solid support, thereby rendering the separation possible.

According to another approach, the single stranded ligation product is obtained by making the template strand double stranded. The double stranded template can be produced by competition hybridisation of a nucleotide similar to the ligation product, or by annealing a primer to the template and extending said primer over the extent of the template using a polymerase.

Various types of reactions are possible between reactant according to the invention. In one aspect, the reaction of the chemical entity of an incorporated building block with a

reactive site is an acylation reaction. Suitably the reactive site is an amine and the bond form is an amide bond. Other types of reactions include alkylating reactions, in which a carbon-carbon single bond is formed, and Wittig type reactions, in which a carbon-carbon double bond is formed.

Linkers between chemical entities and the anticodon may be maintained or cleaved following the reaction of the chemical entities. When more than a single chemical entity is reacted usually one or more bond between the encoded molecule or nascent encoded molecule are cleaved to present the display molecule more efficient to e.g. a target.

In the method depicted above, steps ii) through iv) may be repeated as appropriate using a nascent complex as the template and anticodon(s) directed to a non-used codon in the building blocks to be incorporated. The repetition of the process steps allows for a multi-step incorporation and reaction of building block. Multi-step incorporation may be of advantage because separate reaction conditions may be used for each chemical entity to be reacted with the nascent encoded molecule, thereby allowing for a wider range of possible reactions.

Following the formation of the encoded molecule, a post treatment may be provided for. The post treatment may involve cleavage of bonds such that the encoded molecule is maintained connected to the template through a single bond only. Post treatment may also involve deprotection, i.e. removal of protective groups used during the reactions of the chemical entities.

According to a preferred aspect of the invention, a plurality of templates and building blocks are processed simultaneously or sequentially forming a library of complexes. Suitably, a large pool of templates, such as 108 are provided. This pool of templates is contacted with a pool of building blocks directed to the each codon of the plurality of templates. Preferably two or more pools of building blocks are added sequentially to obtain a multi-step incorporation and reaction. In one aspect of the invention the nucleotide sequences harbouring the different anticodons in each pool have an identical flanking sequence to ensure that the incorporation will occur in frame.

The invention also relates to a library of different complexes, each complex comprising an encoded molecule and a template, which has encoded the chemical entities which has participated in the synthesis thereof, said library being obtainable by processing a plurality of different templates and a plurality of building blocks as depicted above.

The invention also pertains to a method comprising subjecting the library of complexes to a condition partitioning complexes displaying a predetermined property from the remainder of the library. The condition for partitioning of the desired complexes can include subjecting the library of complexes to a molecular target and partitioning complexes binding to said target. Subsequently, nucleic acid sequences comprising the codons and/or the anticodons and/or sequences complementary thereto may be recovered from the partitioned complexes. The nucleic acid sequences of the partitioned complexes are preferably amplified to produce more copies of the templates from successful complexes. In a preferred aspect the nucleic acid sequences of the partitioned complexes are amplified using the polymerase chain reaction (PCR). In one aspect, the amplification product is used to prepare one or more templates which may be utilized in the method of the invention.

Several rounds of synthesis of bifunctional molecules, partitioning of complexes having a desired property, and amplification of templates from complexes having the desired properties can be conducted. As an example, 2 to 15 rounds may applied, suitable 3 to 7 rounds.

Template

It is preferred that the template is divided into coding regions or codons, which codes for specific chemical entities. A codon is a sequence of nucleotides or a single nucleotide. The templates are usually amplifiable and the nucleobases are in a certain aspect selected from the natural nucleobases (adenine, guanine, uracil, thymine, and cytosine) and the backbone is selected from DNA or RNA, preferably DNA.

In the generation of a library, a codon of a single nucleotide will allow for the incorporation of four different chemical entities into the encoded molecule, using the four natural DNA nucleobases (A, C, T, and G). However, to obtain a higher diversity, a codon in certain embodiments preferably comprises at least two and more preferred at least three nucleotides. Theoretically, this will provide for 4^2 and 4^3, respectively, different

chemical entities. The codons will usually not comprise more than 200 nucleotides. It is preferred to have codons with a sequence of 3 to 300 nucleotides, more preferred 4 to 15 nucleotides.

The template sequence will in general have at least two codons which are arranged in sequence, i.e. next to each other. Each of the codons may be separated by a framing sequence. Depending on the encoded molecule formed, the template sequence may comprise further codons, such as 3, 4, 5, or more codons. Each of the further codons may be separated by a suitable framing sequence. Preferably, all or at least a majority of the codons of the nucleic acid sequence are arranged in sequence and each of the codons is separated from a neighbouring codon by a framing sequence. The framing sequence may have any appropriate number of nucleotides, e.g. 1 to 20. Alternatively, codons on the template may be designed with overlapping sequences.

Generally, it is preferred to have more than two codons on the template to allow for the synthesis of more diverse encoded molecules. In a preferred aspect of the invention the number of codons of the template sequence is 2 to 100, more preferred the template sequences comprises 3 to 20 codons.

The framing sequence may serve various purposes. In one setup of the invention, the framing sequence identifies the position of a codon. Usually, the framing sequence either upstream or downstream of a codon comprises information which allows determination of the position of the codon.

The framing sequence may also or in addition provide for a region of high affinity. The high affinity region may ensure that the hybridisation of the template sequence with the anti-codon will occur in frame. Moreover, the framing sequence may adjust the annealing temperature to a desired level.

A framing sequence with high affinity can be provided by incorporation of one or more nucleobases forming three hydrogen bonds to a cognate nucleobase. An example of nucleobases displaying this property is guanine and cytosine. Alternatively, or in addition, the framing sequence may be subjected to back bone modification. Several back bone modifications provides for higher affinity, such as 2'-O-methyl substitution of the ribose moiety, peptide nucleic acids (PNA), and 2'-4' O-methylene cyclisation of the ribose moiety, also referred to as LNA (Locked Nucleic Acid).

The template sequence may comprise flanking regions around the coding segments. The flanking regions can serve as priming sites for an amplification reaction, such as PCR. The template may in certain embodiments comprise a region complementary to the flaking region to allow for a hairpin loop to be formed.

It is to be understood that when the term template is used in the present description and claims, the sequence may be in the sense or the anti-sense format, i.e. the template sequence can be a sequence of codons which actually codes for the molecule or can be a sequence complementary thereto.

It is within the capability of the skilled person in the art to construct the desired design of an oligonucleotide. When a certain annealing temperature is desired it is a standard procedure to suggest appropriate compositions of nucleic acid monomers and the length thereof. The construction of an appropriate design may be assisted by software, such as Vector NTI Suite or the public database at the internet address http://www.nwfsc.noaa.gov/protocols/oligoTMcalc.html.

The conditions which allow hybridisation of the template with a nucleic acid, such as an anti-codon, are influenced by a number of factors including temperature, salt concentration, type of buffer, and acidity. It is within the capabilities of the person skilled in the art to select appropriate conditions to ensure that the contacting between the template sequences and the building blocks are performed at hybridisation conditions. The temperature at which two single stranded oligonucleotides forms a duplex is referred to as the annealing temperature or the melting temperature.

Encoded molecule

The encoded molecule may be formed by a variety of reactants which are reacted with each other and/or a scaffold molecule. Optionally, this reaction product may be post-modified to obtain the final display molecule. The post-modification may involve the cleavage of one or more chemical bonds attaching the encoded molecule to the template in order more efficiently to display the encoded molecule.

The formation of an encoded molecule generally starts by a scaffold, i.e. a chemical unit having one or more reactive groups capable of forming a connection to another

reactive group positioned on a chemical entity, thereby generating an addition to the original scaffold. A second chemical entity may react with a reactive group also appearing on the original scaffold or a reactive group incorporated by the first chemical entity. Further chemical entities may be involved in the formation of the final reaction product. The formation of a connection between the chemical entity and the nascent encoded molecule may be mediated by a bridging molecule. As an example, if the nascent encoded molecule and the chemical entity both comprise an amine group a connection between these can be mediated by a dicarboxylic acid.

The encoded molecule may be attached directly to the template sequence or through a suitable linking moiety. Furthermore, the encoded molecule may be linked to the template sequence through a cleavable linker to release the encoded molecule at a point in time selected by the experimenter.

The chemical entities are suitably mediated to the nascent encoded molecule by a building block, which further comprises an anticodon. The anti-codon serves the function of transferring the genetic information of the building block in conjunction with the transfer of a chemical entity to the nascent complex. The chemical entities are preferably reacted without enzymatic interaction. Notably, the reaction of the chemical entities is preferably not mediated by ribosomes or enzymes having similar activity.

The chemical entity of the building block may in most cases be regarded as a precursor for the structural entity eventually incorporated into the encoded molecule. In other cases the chemical entity provides for the eliminations of chemical units of the nascent scaffold. Therefore, when it in the present application with claims is stated that a chemical entity is transferred to a nascent encoded molecule or a reactive site, it is to be understood that not necessarily all the atoms of the original chemical entity is to be found in the eventually formed encoded molecule. Also, as a consequence of the reactions involved in the connection, the structure of the chemical entity can be changed when it appears on the nascent encoded molecule. Especially, the cleavage resulting in the release of the entity may generate a reactive group which in a subsequent step can participate in the formation of a connection between a nascent complex and a chemical entity.

Building block

The chemical entities that are precursors for structural additions or eliminations of the encoded molecule may be attached to a building block prior to the participation in the formation of the reaction product leading the final encoded molecule. Besides the chemical entity, the building block generally comprises an anti-codon.

The chemical entity of the building block comprises at least one reactive group capable of participating in a reaction which results in a connection between the chemical entity of the building block and another chemical entity or a scaffold associated with the nascent complex. The connection is facilitated by one or more reactive groups of the chemical entity. The number of reactive groups which appear on the chemical entity is suitably one to ten. A building block featuring only one reactive group is used i.a. in the end positions of polymers or scaffolds, whereas building blocks having two reactive groups are suitable for the formation of the body part of a polymer or scaffolds capable of being reacted further. One, two or more reactive groups intended for the formation of connections, are typically present on scaffolds.

The reactive group of the building block may be capable of forming a direct connection to a reactive group of the nascent complex or the reactive group of the building block may be capable of forming a connection to a reactive group of the nascent complex through a bridging fill-in group. It is to be understood that not all the atoms of a reactive group are necessarily maintained in the connection formed. Rather, the reactive groups are to be regarded as precursors for the structure of the connection.

The subsequent cleavage step to release the chemical entity from the building block can be performed in any appropriate way. In an aspect of the invention the cleavage involves usage of a reagent or an enzyme. The cleavage results in a transfer of the chemical entity to the nascent encoded molecule or in a transfer of the nascent encoded molecule to the chemical entity of the building block. In some cases it may be advantageous to introduce new chemical groups as a consequence of linker cleavage. The new chemical groups may be used for further reaction in a subsequent cycle, either directly or after having been activated. In other cases it is desirable that no trace of the linker remains after the cleavage.

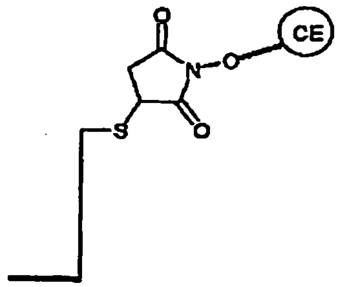
In another aspect, the connection and the cleavage is conducted as a simultaneous reaction, i.e. either the chemical entity of the building block or the nascent encoded molecule is a leaving group of the reaction. In general, it is preferred to design the system such that the connection and the cleavage occur simultaneously because this will reduce the number of steps and the complexity. The simultaneous connection and cleavage can also be designed such that either no trace of the linker remains or such that a new chemical group for further reaction is introduced, as described above.

The attachment of the chemical entity to the building block, optionally via a suitable spacer can be at any entity available for attachment, e.g. the chemical entity can be attached to a nucleobase or the backbone. In general, it is preferred to attach the chemical entity at the phosphor of the internucleoside linkage or at the nucleobase. When the nucleobase is used for attachment of the chemical entity, the attachment point is usually at the 7 position of the purines or 7-deaza-purins or at the 5 position of pyrimidines. The nucleotide may be distanced from the reactive group of the chemical entity by a spacer moiety. The spacer may be designed such that the conformational space sampled by the reactive group is optimized for a reaction with the reactive group of the nascent encoded molecule or reactive site.

The anticodon complements the codon of the template sequence and generally comprises the same number of nucleotides as the codon. The anticodon may be adjoined with a fixed sequence, such as a sequence complementing a framing sequence.

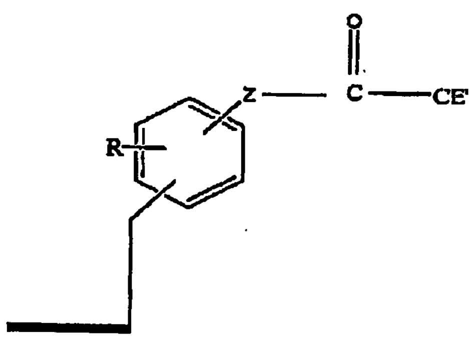
Various specific building blocks are envisaged. Building blocks of particular interest are shown below.

<u>Building blocks transferring a chemical entity to a recipient nucleophilic group</u>
The building block indicated below is capable of transferring a chemical entity (CE) to a recipient nucleophilic group, typically an amine group. The bold lower horizontal line illustrates the building block and the vertical line illustrates a spacer. The 5-membered substituted N-hydroxysuccinimid (NHS) ring serves as an activator, i.e. a labile bond is formed between the oxygen atom connected to the NHS ring and the chemical entity. The labile bond may be cleaved by a nucleophilic group, e.g. positioned on a scaffold



The 5-membered substituted N-hydroxysuccinimid (NHS) ring serves as an activator, i.e. a labile bond is formed between the oxygen atom connected to the NHS ring and the chemical entity. The labile bond may be cleaved by a nucleophilic group, e.g. positioned on a scaffold, to transfer the chemical entity to the scaffold, thus converting the remainder of the fragment into a leaving group of the reaction. When the chemical entity is connected to the activator through an carbonyl group and the recipient group is an amine, the bond formed on the scaffold will an amide bond. The above building block is the subject of the Danish patent application No. PA 2002 01946 and the US provisional patent application No. 60/434,439, the content of which are incorporated herein in their entirety by reference.

Another building block which may form an amide bond is

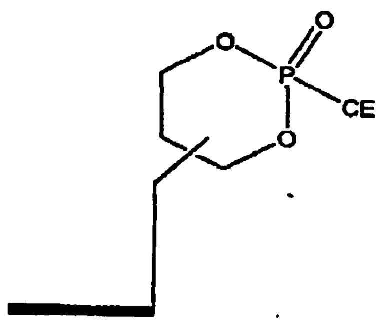


R may be absent or NO_2, CF_3, halogen, preferably Cl, Br, or I, and Z may be S or O. This type of building block is disclosed in Danish patent application No. PA 2002 0951 and US provisional patent application filed 20 December 2002 with the title "A building block capable of transferring a functional entity to a recipient reactive group". The content of both patent application are incorporated herein in their entirety by reference.

A nucleophilic group can cleave the linkage between Z and the carbonyl group thereby transferring the chemical entity –(C=O)-CE' to said nucleophilic group.

Building blocks transferring a chemical entity to a recipient reactive group forming a C=C bond

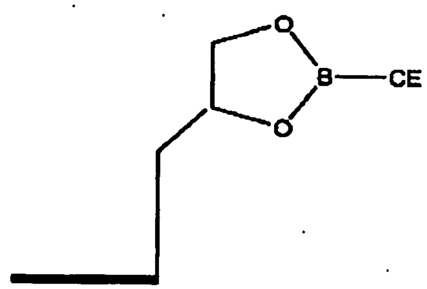
A building block as shown below are able to transfer the chemical entity to a recipient aldehylde group thereby forming a double bond between the carbon of the aldehyde and the chemical entity



The above building block is comprised by the Danish patent application No. DK PA 2002 01952 and the US provisional patent application filed 20 December 2002 with the title "A building block capable of transferring a functional entity to a recipient reactive group forming a C=C double bond". The content of both patent applications are incorporated herein in their entirety by reference.

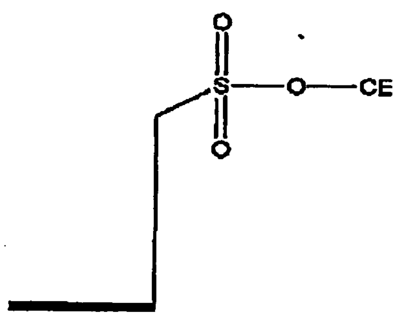
Building blocks transferring a chemical entity to a recipient reactive group forming a C-C bond

The below building block is able to transfer the chemical entity to a recipient group thereby forming a single bond between the receiving moiety, e.g. a scaffold, and the chemical entity.



The above building block is comprised by the Danish patent application No. DK PA 2002 01947 and the US provisional patent application No 60/434,428. The content of both patent applications are incorporated herein in their entirety by reference.

Another building block capable of transferring a chemical entity to a receiving reactive group forming a single bond is



The receiving group may be a nucleophile, such as a group comprising a hetero atom, thereby forming a single bond between the chemical entity and the hetero atom, or the receiving group may be an electronegative carbon atom, thereby forming a C-C bond between the chemical entity and the scaffold.

The chemical entity attached to any of the above building blocks may be a selected from a large arsenal of chemical structures. Examples of chemical entities are H or entities selected among the group consisting of a C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_4-C_8 alkadienyl, C_3-C_7 cycloalkyl, C_3-C_7 cycloheteroalkyl, aryl, and heteroaryl, said group being substituted with 0-3 R^4, 0-3 R^5 and 0-3 R^9 or C_1-C_3 alkylene-NR^4_2, C_1-C_3 alkylene-$NR^4C(O)R^8$, C_1-C_3 alkylene-$NR^4C(O)OR^8$, C_1-C_2 alkylene-O-NR^4_2, C_1-C_2 alkylene-O-$NR^4C(O)R^8$, C_1-C_2 alkylene-O-$NR^4C(O)OR^8$ substituted with 0-3 R^9.

where R^4 is H or selected independently among the group consisting of C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_3-C_7 cycloalkyl, C_3-C_7 cycloheteroalkyl, aryl, heteroaryl, said group being substituted with 0-3 R^9 and

R^5 is selected independently from -N_3, -CNO, -C(NOH)NH_2, -NHOH, -NHNHR^6, -C(O)R^6, -SnR^6_3, -B(OR^6)$_2$, -P(O)(OR^6)$_2$ or the group consisting of C_2-C_8 alkenyl, C_2-C_8 alkynyl, C_4-C_8 alkadienyl said group being substituted with 0-2 R^7,

where R^6 is selected independently from H, C_1-C_6 alkyl, C_3-C_7 cycloalkyl, aryl or C_1-C_6 alkylene-aryl substituted with 0-5 halogen atoms selected from -F, -Cl, -Br, and -I; and R^7 is independently selected from -NO_2, -COOR^6, -COR^6, -CN, -OSiR^6_3, -OR^6 and -NR^6_2.

R^8 is H, C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_3-C_7 cycloalkyl, aryl or C_1-C_6 alkylene-aryl substituted with 0-3 substituents independently selected from -F, -Cl, -NO_2, -R^3, -OR^3, -SiR^3_3

R^9 is =O, -F, -Cl, -Br, -I, -CN, -NO_2, -OR^6, -NR^6_2, -NR^6-C(O)R^6, -NR^6-C(O)OR^6, -SR^6, -S(O)R^6, -S(O)$_2R^6$, -COOR^6, -C(O)NR^6_2 and -S(O)$_2$NR^6_2.

Partitioning

The partition step may be referred to as a selection or a screen, as appropriate, and includes the screening of the library for encoded molecules having predetermined desirable characteristics. Predetermined desirable characteristics can include binding to a target, catalytically changing the target, chemically reacting with a target in a manner which alters/modifies the target or the functional activity of the target, and covalently attaching to the target as in a suicide inhibitor.

The target can be any compound of interest. E.g. the target can be a protein, peptide, carbohydrate, polysaccharide, glycoprotein, hormone, receptor, antigen, antibody, virus, substrate, metabolite, transition state analogue, cofactor, inhibitor, drug, dye, nutrient, growth factor, cell, tissue, etc. without limitation. Particularly preferred targets include, but are not limited to, angiotensin converting enzyme, renin, cyclooxygenase, 5-lipoxygenase, IIL- 1 0 converting enzyme, cytokine receptors, PDGF receptor, type II inosine monophosphate dehydrogenase, β-lactamases, integrin, and fungal cytochrome P-450. Targets can include, but are not limited to, bradykinin, neutrophil elastase, the HIV proteins, including *tat, rev, gag, int,* RT, nucleocapsid etc., VEGF, bFGF, TGFβ, KGF, PDGF, thrombin, theophylline, caffeine, substance P, IgE, sPLA2, red

blood cells, glioblastomas, fibrin clots, PBMCs, hCG, lectins, selectins, cytokines, ICP4, complement proteins, etc.

Encoded molecules having predetermined desirable characteristics can be partitioned away from the rest of the library while still attached to the template sequence by various methods known to one of ordinary skill in the art. In one embodiment of the invention the desirable products are partitioned away from the entire library without chemical degradation of the attached nucleic acid template such that the templates are amplifiable. The templates may then be amplified, either still attached to the desirable encoded molecule or after separation from the desirable encoded molecule.

In a preferred embodiment, the desirable encoded molecule acts on the target without any interaction between the nucleic acid attached to the desirable encoded molecule and the target. In one embodiment, the bound complex-target aggregate can be partitioned from unbound complexes by a number of methods. The methods include nitrocellulose filter binding, column chromatography, filtration, affinity chromatography, centrifugation, and other well known methods.

Briefly, the library of complexes is subjected to the partitioning step, which may include contact between the library and a column onto which the target is immobilised. Templates associated with undesirable encoded molecules, i.e. encoded molecules not bound to the target under the stringency conditions used, will pass through the column. Additional undesirable encoded molecules (e.g. encoded molecules which cross-react with other targets) may be removed by counter-selection methods. Desirable complexes are bound to the column and can be eluted by changing the conditions of the column (e.g., salt, pH, surfactant, etc.) or the template.

Additionally, chemical compounds which react with a target can be separated from those products that do not react with the target. In one example, a chemical compound which covalently attaches to the target (such as a suicide inhibitor) can be washed under very stringent conditions. The resulting complex can then be treated with proteinase, DNAse or other suitable reagents to cleave a linker and liberate the nucleic acids which are associated with the desirable chemical compound. The liberated nucleic acids can be amplified.

In another example, the predetermined characteristic of the desirable product is the ability of the product to transfer a chemical group (such as acyl transfer) to the target and thereby inactivate the target. One could have a product library where all of the products have a thioester chemical group. Upon contact with the target, the desirable products will transfer the chemical group to the target concomitantly changing the desirable product from a thioester to a thiol. Therefore, a partitioning method which would identify products that are now thiols (rather than thioesters) will enable the selection of the desirable products and amplification of the nucleic acid associated therewith.

There are other partitioning and screening processes which are compatible with this invention that are known to one of ordinary skill in the art. In one embodiment, the products can be fractionated by a number of common methods and then each fraction is then assayed for activity. The fractionization methods can include size, pH, hydrophobicity, etc.

Inherent in the present method is the selection of encoded molecules on the basis of a desired function; this can be extended to the selection of molecules with a desired function and specificity. Specificity can be required during the selection process by first extracting template sequences of chemical compounds which are capable of interacting with a non-desired "target" (negative selection, or counter-selection), followed by positive selection with the desired target. As an example, inhibitors of fungal cytochrome P-450 are known to cross-react to some extent with mammalian cytochrome P-450 (resulting in serious side effects). Highly specific inhibitors of the fungal cytochrome could be selected from a library by first removing those products capable of interacting with the mammalian cytochrome, followed by retention of the remaining products which are capable of interacting with the fungal cytochrome.

Determining the template sequence

The nucleotide sequence of the template sequence present in the isolated bifunctional molecules is determined to identify the chemical entities that participated in the preselected binding interaction.

Although conventional DNA sequencing methods are readily available and useful for

this determination, the amount and quality of isolated bifunctional molecule may require additional manipulations prior to a sequencing reaction.

Where the amount is low, it is preferred to increase the amount of the template sequence by polymerase chain reaction (PCR) using PCR primers directed primer binding sites present in the template sequence.

In addition, the quality of the isolated bifunctional molecule may be such that multiple species of bifunctional molecule are co-isolated by virtue of similar capacities for binding to the target. In cases where more than one species of bifunctional molecule are isolated, the different isolated species must be separated prior to sequencing of the identifier oligonucleotide.

Thus in one embodiment, the different template sequences of the isolated bifunctional complexes are cloned into separate sequencing vectors prior to determining their sequence by DNA sequencing methods. This is typically accomplished by amplifying all of the different template sequences by PCR as described herein, and then using a unique restriction endonuclease sites on the amplified product to directionally clone the amplified fragments into sequencing vectors. The cloning and sequencing of the amplified fragments then is a routine procedure that can be carried out by any of a number of molecular biological methods known in the art.

Alternatively, the bifunctional complex or the PCR amplified template sequence can be analysed in a microarray. The array may be designed to analyse the presence of a single codon or multiple codons in a template sequence.

Nucleotides

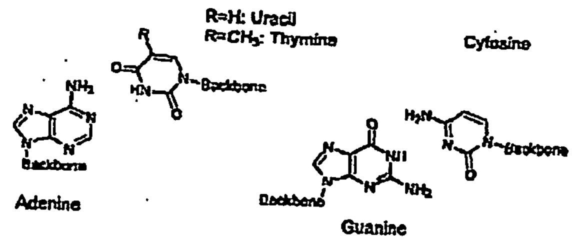
The nucleic acids used in the present invention may be a single nucleotide or several nucleotides linked together in an oligonucleotide. Each nucleotide monomer is normally composed of two parts, namely a nucleobase moiety, and a backbone. The back bone may in some cases be subdivided into a sugar moiety and an internucleoside linker.

The nucleobase moiety may be selected among naturally occurring nucleobases as well as non-naturally occurring nucleobases. Thus, "nucleobase" includes not only the

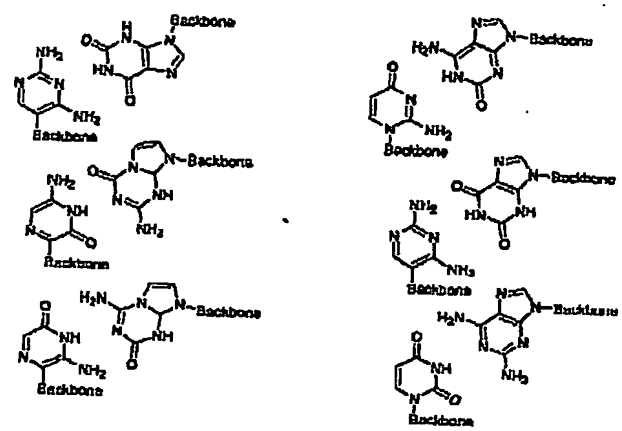
known purine and pyrimidine hetero-cycles, but also heterocyclic analogues and tautomers thereof. Illustrative examples of nucleobases are adenine, guanine, thymine, cytosine, uracil, purine, xanthine, diaminopurine, 8-oxo-N^6-methyladenine, 7-deazaxanthine, 7-deazaguanine, N^4,N^4-ethanocytosin, N^6,N^6-ethano-2,6-diamino-purine, 5-methylcytosine, 5-(C^3-C^6)-alkynylcytosine, 5-fluorouracil, 5-bromouracil, pseudoisocytosine, 2-hydroxy-5-methyl-4-triazolopyridine, isocytosine, isoguanine, inosine and the "non-naturally occurring" nucleobases described in Benner et al., U.S. Pat No. 5,432,272. The term "nucleobase" is intended to cover these examples as well as analogues and tautomers thereof. Especially interesting nucleobases are adenine, guanine, thymine, cytosine, 5-methylcytosine, and uracil, which are considered as the naturally occurring nucleobases.

Examples of suitable specific pairs of nucleobases are shown below:

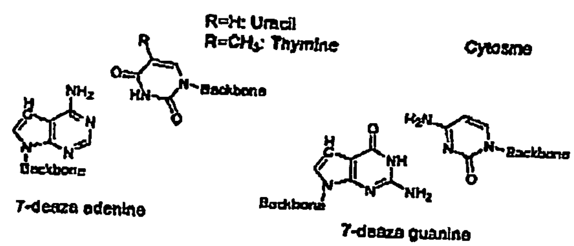
Natural Base Pairs



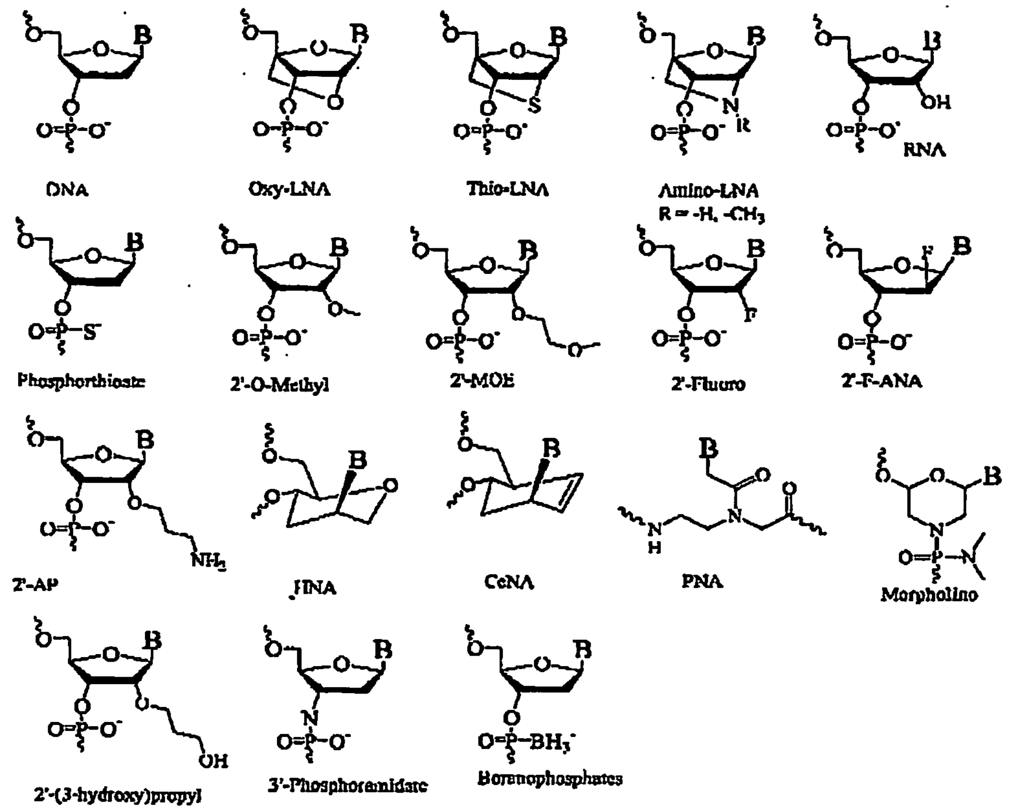
Synthetic Base Pairs



Synthetic purine bases pairing with natural pyrimidines



Suitable examples of backbone units are shown below (B denotes a nucleobase):



The sugar moiety of the backbone is suitably a pentose but may be the appropriate part of an PNA or a six-member ring. Suitable examples of possible pentoses include ribose, 2'-deoxyribose, 2'-O-methyl-ribose, 2'-flour-ribose, threonucleic acid (TNA), and 2'-4'-O-methylene-ribose (LNA). Suitably the nucleobase is attached to the 1' position of the pentose entity.

An internucleoside linker connects the 3' end of preceding monomer to a 5' end of a succeeding monomer when the sugar moiety of the backbone is a pentose, like ribose or 2-deoxyribose. The internucleoside linkage may be the natural occurring phospodiester linkage or a derivative thereof. Examples of such derivatives include phosphorothioate, methylphosphonate, phosphoramidate, phosphotriester, and phosphodithioate. Furthermore, the internucleoside linker can be any of a number of non-phosphorous-containing linkers known in the art.

Preferred nucleic acid monomers include naturally occurring nucleosides forming part of the DNA as well as the RNA family connected through phosphodiester linkages. The members of the DNA family include deoxyadenosine, deoxyguanosine, deoxythymidine, and deoxycytidine. The members of the RNA family include adenosine, guanosine, uridine, cytidine, and inosine. Inosine is a non-specific pairing nucleoside and may be used as universal base because inosine can pair nearly isoenergetically with A, T, and C.

Synthesis of nucleic acids

Oligonucleotides can be synthesized by a variety of chemistries as is well known. For synthesis of an oligonucleotide on a substrate in the direction of 3' to 5', a free hydroxy terminus is required that can be conveniently blocked and deblocked as needed. A preferred hydroxy terminus blocking group is a dimexothytrityl ether (DMT). DMT blocked termini are first deblocked, such as by treatment with 3% dichloroacetic acid in dichloromethane (DCM) as is well known for oligonucleotide synthesis, to form a free hydroxy terminus.

Nucleotides in precursor form for addition to a free hydroxy terminus in the direction of 3' to 5' require a phosphoramidate moiety having an aminodiisopropyl side chain at the 3' terminus of a nucleotide. In addition, the free hydroxy of the phosphoramidate is blocked with a cyanoethyl ester (OCNET), and the 5' terminus is blocked with a DMT ether. The addition of a 5' DMT-, 3' OCNET-blocked phosphoramidate nucleotide to a free hydroxyl requires tetrazole in acetonitrile followed by iodine oxidation and capping of unreacted hydroxyls with acetic anhydride, as is well known for oligonucleotide synthesis. The resulting product contains an added nucleotide residue with a DMT blocked 5' terminus, ready for deblocking and addition of a subsequent blocked nucleotide as before.

For synthesis of an oligonucleotide in the direction of 5' to 3', a free hydroxy terminus on the linker is required as before. However, the blocked nucleotide to be added has the blocking chemistries reversed on its 5' and 3' termini to facilitate addition in the opposite orientation. A nucleotide with a free 3' hydroxyl and 5' DMT ether is first blocked at the 3' hydroxy terminus by reaction with TBS-Cl in imidazole to form a TBS ester at the 3' terminus. Then the DMT-blocked 5' terminus is deblocked with DCA in DCM as

before to form a free 5' hydroxy terminus. The reagent (N,N-diisopropylamino)(cyanoethyl) phosphonamidic chloride having an aminodiisopropyl group and an OCNET ester is reacted in tetrahydrofuran (THF) with the 5' deblocked nucleotide to form the aminodiisopropyl-, OCNET-blocked phosphonamidate group on the 5' terminus. Thereafter the 3' TBS ester is removed with tetrabutylammonium fluoride (TBAF) in DCM to form a nucleotide with the phosphonamidate-blocked 5' terminus and a free 3' hydroxy terminus. Reaction in base with DMT-Cl adds a DMT ether blocking group to the 3' hydroxy terminus.

The addition of the 3' DMT-, 5' OCNET-blocked phosphonamidated nucleotide to a linker substrate having a free hydroxy terminus then proceeds using the previous tetrazole reaction, as is well known for oligonucleotide polymerization. The resulting product contains an added nucleotide residue with a DMT-blocked 3' terminus, ready for deblocking with DCA in DCM and the addition of a subsequent blocked nucleotide as before.

Extension and amplification

The use of the polymerase chain reaction (PCR) is a preferred embodiment, for the production of the templates using the nucleic acids of the selected complexes as templates.

For use in this invention, the template sequences are preferably comprised of polynucleotide coding strands, such as mRNA and/or the sense strand of genomic DNA or non-natural nucleic acids, like TNA and LNA which may be used as template for a polymerase. If the genetic material to be processed is in the form of double stranded nucleic acid, it is usually first denatured, typically by melting, into single strands. The nucleic acid is subjected to a PCR reaction by treating (contacting) the sample with a PCR primer pair, each member of the pair having a preselected nucleotide sequence. The PCR primer pair is capable of initiating primer extension reactions by hybridizing to the PCR primer binding site on template oligonucleotide, preferably at least about 10 nucleotides in length, more preferably at least about 12 nucleotides in length. The first primer of a PCR primer pair is sometimes referred to as the "anti-sense primer" because it is extended into a non-coding or anti-sense strand of a nucleic acid, i.e., a strand complementary to a coding strand. The second primer of a PCR primer pair is

sometimes referred to as the "sense primer" because it is adjoined with the coding or sense strand of a nucleic acid.

The PCR reaction is performed by mixing the PCR primer pair, preferably a predetermined amount thereof, with the nucleic acids of the sample, preferably a predetermined amount thereof, in a PCR buffer to form a PCR reaction admixture. The admixture is thermocycled for a number of cycles, which is typically predetermined, sufficient for the formation of a PCR reaction product, thereby amplifiyng the templates in the isolated complex.

PCR is typically carried out by thermocycling i.e., repeatedly increasing and decreasing the temperature of a PCR reaction admixture within a temperature range whose lower limit is about 30 degrees Celsius (30° C.) to about 55° C. and whose upper limit is about 90° C. to about 100° C. The increasing and decreasing can be continuous, but is preferably phasic with time periods of relative temperature stability at each of temperatures favoring polynucleotide synthesis, denaturation and hybridization.

A plurality of first primer and/or a plurality of second primers can be used in each amplification, e.g., one species of first primer can be paired with a number of different second primers to form several different primer pairs. Alternatively, an individual pair of first and second primers can be used. In any case, the amplification products of amplifications using the same or different combinations of first and second primers can be combined for assaying for mutations.

The PCR reaction is performed using any suitable method. Generally it occurs in a buffered aqueous solution, i.e., a PCR buffer, preferably at a pH of 7-9, most preferably about 8. Preferably, a molar excess of the primer is admixed to the buffer containing the template strand. A large molar excess is preferred to improve the efficiency of the process.

The PCR buffer also contains the deoxyribonucleotide triphosphates (polynucleotide synthesis substrates) dATP, dCTP, dGTP, and dTTP and a polymerase, typically thermostable, all in adequate amounts for primer extension (polynucleotide synthesis) reaction. The resulting solution (PCR admixture) is heated to about 90° C.-100° C. for about 1 to 10 minutes, preferably from 1 to 4 minutes. After this heating period the so-

lution is allowed to cool to a primer hybridization temperature. The synthesis reaction may occur at from room temperature up to a temperature above which the polymerase (inducing agent) no longer functions efficiently. Thus, for example, if DNA polymerase is used as inducing agent, the temperature is generally no greater than about 40° C. The thermocycling is repeated until the desired amount of PCR product is produced. An exemplary PCR buffer comprises the following: 50 mM KCl; 10 mM Tris-HCl at pH 8.3; 1.5 mM MgCl2 ; 0.001% (wt/vol) gelatin, 200 μM dATP; 200 μM dTTP; 200 μM dCTP; 200 μM dGTP; and 2.5 units Thermus aquaticus (Taq) DNA polymerase I (U.S. Pat. No. 4,889,818) per 100 microliters (μl) of buffer.

The inducing agent may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, E. coli DNA polymerase I, Klenow fragment of E. coli DNA polymerase I, T4 DNA polymerase, Taq DNA polymerase, Pfu polymerase, Vent polymerase, HIV-1 Reverse Transcriptase, other available DNA polymerases, reverse transcriptase, and other enzymes, including heat-stable enzymes, which will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each nucleic acid strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the template strand, until synthesis terminates, producing molecules of different lengths. There may be inducing agents, however, which initiate synthesis at the 5' end and proceed in the above direction, using the same process as described above.

The inducing agent also may be a compound or system which will function to accomplish the synthesis of RNA primer extension products, including enzymes. In preferred embodiments, the inducing agent may be a DNA-dependent RNA polymerase such as T7 RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase. These polymerases produce a complementary RNA polynucleotide. The high turn-over rate of the RNA polymerase amplifies the starting polynucleotide as has been described by Chamberlin et al., The Enzymes, ed. P. Boyer, pp. 87-108, Academic Press, New York (1982). Amplification systems based on transcription have been described by Gingeras et al., in PCR Protocols, A Guide to Methods and Applications, pp. 245-252, Innis et al., eds, Academic Press, Inc., San Diego, Calif. (1990).

If the inducing agent is a DNA-dependent RNA polymerase and, therefore incorporates ribonucleotide triphosphates, sufficient amounts of ATP, CTP, GTP and UTP are admixed to the primer extension reaction admixture and the resulting solution is treated as described above.

The newly synthesized strand and its complementary nucleic acid strand form a double-stranded molecule which can be used in the succeeding steps of the method. PCR amplification methods are described in detail in U.S. Pat. Nos. 4,683,192, 4,683,202, 4,800,159, and 4,965,188, and at least in several texts including PCR Technology: Principles and Applications for DNA Amplification, H. Erlich, ed., Stockton Press, New York (1989); and PCR Protocols: A Guide to Methods and Applications, Innis et al., eds., Academic Press, San Diego, Calif. (1990). The term "primer" as used herein refers to a polynucleotide whether purified from a nucleic acid restriction digest or produced synthetically, which is capable of acting as a point of initiation of nucleic acid synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, i.e., in the presence of nucleotides and an agent for polymerization such as DNA polymerase, reverse transcriptase and the like, and at a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency, but may alternatively be in double stranded form. If double stranded, the primer is first treated to separate it from its complementary strand before being used to prepare extension products. Preferably, the primer is a polydeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the agents for polymerization. The exact lengths of the primers will depend on many factors, including temperature and the source of primer. For example, depending on the complexity of the target sequence, a polynucleotide primer typically contains 10 to 25 or more nucleotides, although it can contain fewer nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with template.

The primers used herein are selected to be "substantially" complementary to the different strands of each specific sequence to be synthesized or amplified. This means that the primer must be sufficiently complementary to non-randomly hybridize with its respective template strand. Therefore, the primer sequence may or may not reflect the exact sequence of the template. For example, a non-complementary nucleic acid can be attached to the 5' end of the primer, with the remainder of the primer sequence be-

ing substantially complementary to the strand. Such non-complementary fragments typically code for an endonuclease restriction site or used as a linker to connect to a label, such as biotin.

Primers of the present invention may also contain a DNA-dependent RNA polymerase promoter sequence or its complement. See for example, Krieg et al., Nucl. Acids Res., 12:7057-70 (1984); Studier et al., J. Mol. Biol., 189:113-130 (1986); and Molecular Cloning: A Laboratory Manual, Second Edition, Maniatis et al., eds., Cold Spring Harbor, N.Y. (1989).

When a primer containing a DNA-dependent RNA polymerase promoter is used, the primer is hybridized to the polynucleotide strand to be amplified and the second polynucleotide strand of the DNA-dependent RNA polymerase promoter is completed using an inducing agent such as E. coli DNA polymerase I, or the Klenow fragment of E. coli DNA polymerase. The starting polynucleotide is amplified by alternating between the production of an RNA polynucleotide and DNA polynucleotide. This may be used for selective degradation of the RNA strand, which is prone to disintegration upon treatment with a strong base.

Primers may also contain a template sequence or replication initiation site for a RNA-directed RNA polymerase. Typical RNA-directed RNA polymerase include the QB replicase described by Lizardi et al., Biotechnology, 6:1197-1202 (1988). RNA-directed polymerases produce large numbers of RNA strands from a small number of template RNA strands that contain a template sequence or replication initiation site. These polymerases typically give a one million-fold amplification of the template strand as has been described by Kramer et al., J. Mol. Biol., 89:719-736 (1974).

In one embodiment, the present invention utilizes a set of polynucleotides that form primers having a priming region located at the 3'-terminus of the primer. The 3'-terminal priming portion of each primer is capable of acting as a primer to catalyze nucleic acid synthesis, i.e., initiate a primer extension reaction off its 3' terminus. One or both of the primers can additionally contain a 5'-terminal non-priming portion, i.e., a region that does not participate in hybridization to the preferred template. The 5'-part of the primer may be labelled as described herein above.

Brief Description of the Figures

Fig. 1 discloses a general method for producing an encoded molecule using stepwise ligation and stepwise reaction of chemical entities.

Fig. 2 shows a general method for single-step ligation of multiple building blocks.

Fig. 3 shows an oligo-architecture alternating reading frame determinants are used for stepwise ligation of building blocks.

Fig. 4 discloses a photograph of a gel mentioned in example 2.

Fig. 5 discloses a reaction scheme in which a solid support is used for carrying the building block, and

Fig. 6 shows a schematic representation of a set-up useful for the stepwise ligation of anticodon.

Fig. 7 shows a synthesis scheme for the reactions used in example 3.

Detailed Description of the Invention

Fig. 1 discloses the principles of stepwise ligation and stepwise reaction. Initially, a template comprising a hairpin loop is provided. The hairpin loop is formed due to the fact that an outer sequence, such as a flanking sequence is complementary to a sequence in the interior of the sequence harbouring the template. Under hybridisation conditions the complementary sequences will anneal to each other thus forming a starting point for a ligase at the one end of the oligonucleotide. In a subsequent step a building block is added. The building block comprises a nucleic acid sequence complementary to the sequence next to the interior sequence. Either of the ends of the abutting nucleotides generally comprises a phosphate group to make it possible for a ligase to perform the action of ligating the ends together, thereby forming a contigues nucleotide sequence.

The amount of building block added is generally in excess to ensure sufficient substrate for the ligase and a complete as possible reaction. After the ligation step the excess building block not ligated to the template-primer complex is removed and a second building block is added. The second building block has an anticodon complementary to a sequence of the template such that one end of the anticodon of the second building block abuts the anticodon of the preceding incorporated building block. Under suitable hybridisation conditions a ligase is added to ligate the second anticodon to the preceding ligation product to form a single nucleotide sequence comprising the tem-

plate and the two building blocks. Subsequent to the ligation, excess amounts of the second building block is removed.

The reactions between the chemical entities can in general not take place when the strand comprising the ligation product is annealed to the template strand because the linker connecting the anticodon and the functional entity is to short for two adjacent chemical entities to be in sufficient proximity for a reaction to occur. Therefore, the reaction is not possible before the ligated strand is single-stranded. The single stranded ligation product is obtained in the present instance by denature the double helix formed by the ligation product and the template. Under denaturing conditions, the chemical entities are reacted, such that one of the chemical entities is transferred to the other.

The process may be repeated an appropriate number of times until a predetermined number and type of chemical entities have reacted to form the final encoded reaction product. Each cycle starts with the addition of a building block, which has an anticodon with a sequence that anneals next to a preceding anticodon. Excess amounts of the building block is subsequently removed and a ligase is added in order to adjoin the building block to the previously formed ligation product. After the ligation, the chemical entity of the just incorporated building block is reacted with the nascent encoded molecule. The formed encoded molecule may be subjected to various alterations, such as linker cleavage, deprotection, intra-molecular reactions, etc to form the final display molecule. The bifunctional complex comprising the display molecule and the ligation product may be subjected to a partition step, as described herein, to select one or more display molecules displaying desired properties.

In some embodiments of the present invention it may be desirable to synthesise polymers or linear molecules using stepwise chemical reactions i.e. ligation and subsequent reaction of a chemical entity before addition of the next building block. This is particularly desirable when each building block is to be assembled into a polymer or linear molecule using identical reaction types. One example of this reaction type is the use of acylation reactions in the formation of amide bonds in the encoded polymers or linear molecules. Thus, a stepwise procedure will prevent the reaction of chemical entities in random order and instead assure that the chemical entities are added in a ordered fashion according to the template sequence. If multiple anticodons were to be ligated in one single step and the attached chemical entities could react at random a

laborious deconvolution step would be required to identify the exact molecules with desired properties. In contrast, if linear or branched molecules is to be synthesised from orthogonal and compatible chemical reactions it may be desirable to ligate several or all anticodons in a single step and subsequent react one or more of the chemical entities.

Fig. 2 discloses a method in which multiple building blocks are annealed together in a single process step. In some embodiments, it may be desirable to ligate two or more anticodons of building blocks to the template-primer complex in a single step. This could be particular useful for the synthesis of molecules when using orthogonal chemistries for the assembly of the encoded molecule.

Initially, a template comprising a hairpin loop is provided. The hairpin loop is formed due to the fact that an outer sequence, such as a flanking sequence is complementary to a sequence in the interior of the sequence harbouring the template. Under hybridisation conditions the complementary sequences will anneal to each other thus forming a starting point for a ligase at the one end of the oligonucleotide. Various building blocks are added subsequently. The anticodons are designed such that they aligns on the template under hybridisation conditions. The alignment is directed by the sequence of the template. Subsequently or simultaneously with the alignment process the anticodons are ligated together by a ligase or similar ligation means. Once the anticodons have been ligated together, the ligation product is made single stranded by inducing denaturing conditions, that is, conditions ensuring that the double helix formed by the ligation product and the template is disrupted. The chemical entities attached to the ligation product, which is in a single stranded state, are reacted all together or at least the majority of the chemical entities is reacted to form a reaction product. The reaction product may be modified by cleavage of one or more linkers connecting the reaction product with the ligation product to display the encoded reaction product more efficiently. It may also be advantageously to cleave at least some of the linkers in order for the ligated strand to anneal to the template before the bifunctional molecule is subjected to a partition process because a single stranded nucleic acid may be affect a the partition process. Especially, single stranded RNA are known to be able to interact with biological molecules. Furthermore, a double stranded nucleic acid is generally more stable compared to a corresponding single stranded molecule, i.e. a double stranded nucleic acid is generally able to withstand higher temperatures and extreme pH values.

A simple method to assure stepwise ligation is to include reading frame determinants in the template sequence, as shown in Fig. 6. Reading frame determinants can be fixed sequences that prevent multiple ligation reactions on the same template. Thus, by alternating the reading frame determinants in the template. only a certain subtype of anticodons can be ligated. In this set-up a primer-template is designed containing 4 codons of 7 nucleotides. The 1st 7 nucleotide codon sequence adjacent to the free 3' OH-group of the template-primer complex is composed of the general sequence: 5'-TNNANNA-3' encoding up to 256 different chemical entities. The 2nd codon is composed of the general sequence: 5'-ANNANNT-3'. The 3rd and 4th codons are identical to the 1st and 2nd codon regions, respectively. Two sets of anticodon–building blocks are prepared where the 1st anticodon set is composed of the general sequence: 5'-pTNNXNNA-3' and the 2nd of of general sequence: 5'-pANNXNNT-3', where p = phosphate and X is Carboxy-dT useful for the attachment of chemical entities. Stepwise ligation is accomplished by ligation of the 1st anticodon set to the template-primer complex shown schematically in fig. 6. Since 1st anticodon set is complementary to codon 1 and 3 only, a ligation reaction using this anticodon subset will result in the sequence-specific ligation of anticodon-building blocks complementary to codon 1. Subsequently, all unused 1st building blocks are removed followed by addition of the second set of building blocks to the template-primer complex containing a ligated 1st building block. This ligation step will result in the sequence-specific addition of an anticodon complementary to codon 2. Excess anticodon-building block 2 is removed. Next, the first set of building blocks are added to ligate an anticodon-building block complementary to the 3rd codon with the template-primer complex carrying anticodon building blocks complementary to codon 1 and 2. After removal of excess anticodon-building blocks, anticodon-building blocks from the 2nd subset are used for ligation at codon position 4. These alternating ligation steps using two or more subsets of anticodon building blocks allows the experimenter to allocate and ligate specific subsets of anticodon-building blocks to specific positions on the template/primer complex using reading frame determinants.

The chemical reactions required to cross-link and/or transfer a building block can be conducted at any time during this protocol. In one embodiment of this invention it may be desirable to cross-link and transfer a building block after each ligation step to assign a specific building block to a specific position in e.g a polymer or linear molecule. If

such a scheme is used a deconvolution procedure will most likely not be required. In some embodiments it may be desirable to have several or all building blocks ligated to the template-primer complex before performing transfer of building blocks. Furthermore, the building blocks may be reacted with one or more other building block or with a functionality inserted in the template-primer complex.

Using the set-up shown in figure 6 it is possible to generate a library of molecules comprising 256^4 = more than a billion different compounds each encoded by a nucleic acid template. Consequently, it is possible to select the compounds having desired properties such as binding to a target (e.g. a receptor protein) or a catalytic function. Following selection, the template(s) that encode the compounds having desired properties is amplified (f.ex by PCR) and used for additional rounds of templated synthesis, selection and amplification. Finally, cloning and sequencing or other means of sequence detection such as sequence-specific array-detection of the selected templates will determine the composition of the selected small molecules.

Examples

Example 1: Ligation of building blocks to template-primer complex

In this set-up a DNA template that allows for hairpin structure formation is used as both a template and primer for the ligation of building block oligonucleotides (Fig 3). Fig. 3 shows template-primer set-up and the complementary 3 oligonucleotide building blocks. The carboxy dT functionality on each building block oligonucleotide is easily introduced at any position in an oligonucleotide using standard phosphoramidite chemistry and can be used as molecular handle for the addition of chemical entities to each specific oligonucleotide sequence.

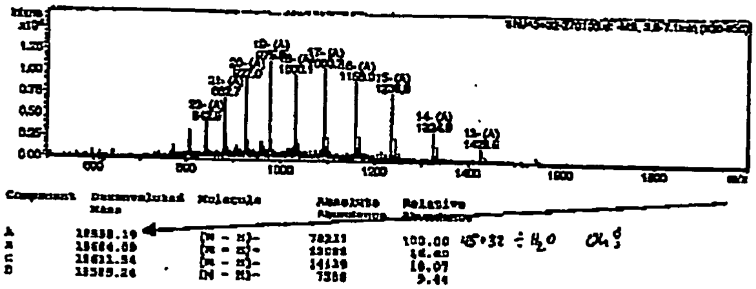
Ligation of building block 1 (BB 1) to the template-primer complex.

500 pmol of 7-mer building block oligonucleotide 1 comprising the carboxy-group was ligated to 500 pmol template-primer complex in a 50 µl volume at 20 °C for 1 hour using Takara ligase kit version 2.0 (TaKaRa™). The sample was extracted twice with 100 µl phenol and purified using double gel-filtration (Biorad microspin 6 columns, Biorad™) and the ligation product was examined by Electrospray MS analysis (ES-MS, Bruker).

BB 1

ACAXCATGGAGCTGGTG T T
AGCAGTTAGTAGTTTGTACTATGTAGTACCTCGACCAC T T T T

Expected mass:
18556 Da



Observed mass:
18558 Da

ES-MS of ligation products obtained from the ligation of template–primer complex and BB1 oligonucleotide shows that the ligation reaction yields a single dominant ligation product corresponding to the template-primer complex ligated to BB1.

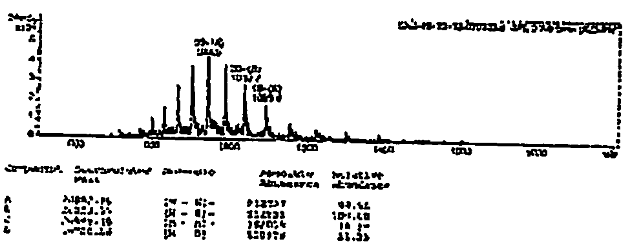
Ligation of BB1 & BB2 to the template-primer complex.

Ligation of equimolar BB1, BB2 and template-primer complex was conducted as described above and the products purified and assayed using ES-MS.

BB 2 BB 1

ACAXGATACAXCATGGAGCTGGTG T T
AGCAGTTAGTAGTTTGTACTATGTAGTACCTCGACCAC T T T T

Expected mass:
20778 Da



Observed mass:
20780 Da.

ES-MS of the BB1, BB2 and template-primer ligation products shows a single dominant product (20802,20823 & 20845 Da species are the mono,- di, tri-sodium complexes of he 20780 Da mass, respectively).

Ligation of BB1, BB2 & BB3 to the template-primer complex.

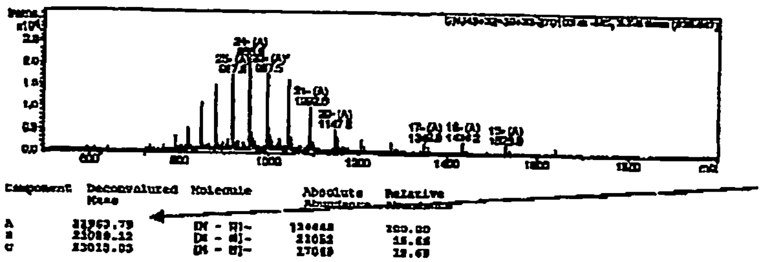
Equimolar BB1, BB2, BB3 and template-primer complex was ligated as described above and examined using ES-MS.

BB 3 BB 2 BB 1

TCAXCAAACAXGATACAXCATGGAGCTGGTG T T

AGCAGTTAGTAGTTTGTACTATGTAGTACCTCGACCAC T T T T

Expected mass: 22961 Da



Observed mass: 22983 Da

ES-MS of BB1, BB2 and BB3 ligated to the template-primer complex shows sequence specific ligation of BB1, BB2, and BB3 oligonucleotides each comprising a central carboxy-dT functionality.

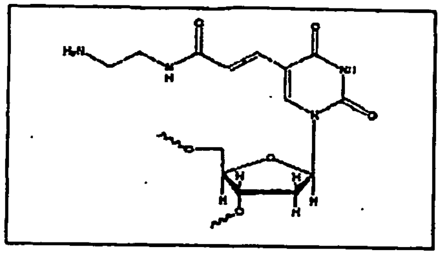
Example 2:

Ligation of multiple building block 6-mer or 12-mer oligonucleotides comprising chemical entities.

In the following example, short 6 or 12-mer oligonucleotides are ligated to flanking oligonucleotides in a templated reaction and the products analysed by polyacrylamide gel electrophoresis (PAGE). Position 3 of each BB-oligonucleotide contains a C2 amino dT which can function as handle for the attachment of desired chemical entities (see scheme below). In this example the BB-oligonucleotides are modified with 4-PBA or SA using the following procedure.

4-PBA (4-penteneoyl-β-alanine) and SA (succinic acid) were both dissolved in 25 µl di-methyl fluoride (DMF), 200 mM final, and mixed with 25 µl 200 mM EDC dissolved in DMF. This mixture was incubated for 30 min at 25°C. Subsequently, 8 nano mole of BB-oligonucleotide dissolved in 50 µl 100 mM Hepes, pH 7.5, was included and allowed to react for 20 min at 25°C. BB-oligonucleotides were extracted twice with ethyl acetate in order to remove unreacted 4-BPA and SA.

C2 Amino-modified dT

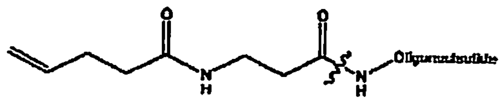


6 mer oligo

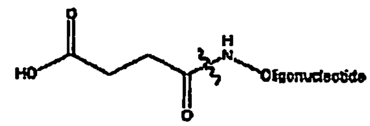
NNN(C2-amino-dT)NN

Loading of Chemical compound (BB)

BB loaded 6 mer oligo



4-penteneoyl-β-alanine (4-PBA) loaded on oligonucleotide



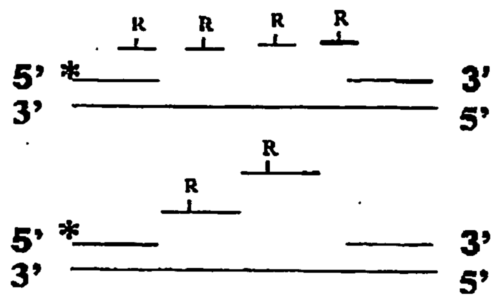
Succinic acid (SA) loaded on oligonucleotide

Structure of C2 amino dT and the chemical modifications 4-PBA and SA loaded on either a 6-mer or 12-mer building block oligonucleotide containing C2 amino dT are shown above.

Modifications on oligonucleotides were verified by ES-MS before subjecting the oligonucleotides to template specific ligation.

The 5' flanking oligo was 5' ^{32}P-labelled using T4 polynucleotide kinase. This was done to be able to visualise mobility shift of ligated products. Two pico mole of template and two pico mole of each of the two flanking oligos were mixed with 20 pico mole of 6-mer or 12-mer BB-oligonucleotides loaded with either 4-BPA or SA, heated 1 minute at 80°C and allowed to anneal by cooling to 20°C. Annealed oligos were ligated at 20°C for 1 hour using Takara ligase kit version 2.0 (TaKaRa™). The samples were denatured in SDS sample buffer and analysed by SDS-PAGE containing 8M UREA.

Schematic representation of experiment



The * marks represents a ^{32}P-radiactively labelled primer and R marks the chemical modification on 6 or 12-mer BB-oligonucleotides. Fig. 4 shows the results of the experiment represented by a gel analysis of the ligation products. In fig. 3 the lanes are numbered:

Lane 1: ^{32}P primer
2: ^{32}P template
3: 4 X 6 mer
4: 2 X 12 mer
5: 4 X 6 mer + 4-PBA
6: 2 X 12 mer + 4PBA
7: 4 X 6 mer + SA
8: 2 X 12 mer + SA

The data shows that multiple 6 or 12-mer building block oligonucleotides can be efficiently ligated to flanking primers in a templated reaction using either 4-PBA or SA as attached chemical entity. Thus, ligation is template specific and unaffected by the attached chemical moiety.

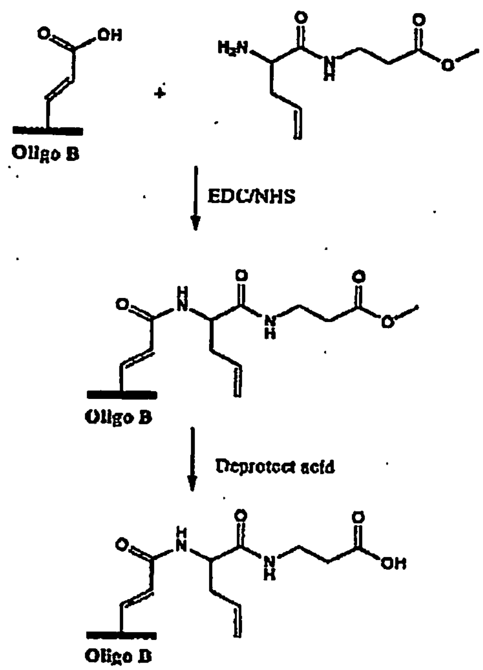
Example 3

Transfer of a functional entity to a reactive site.

A 7-nucleotide anticodon oligonucleotide (A) with the sequence pATGXCAT, where X is C6 amino-dT (Glen Research cat# 10-1039-90) and p is photoprotected 5'-terminal phosphate (Glen Research cat# 10-4913-90) was obtained from DNA technology A/S Denmark. A second 7-nucleotide anticodon oligonucleotide (B) with the sequence 5'-pATCYGTA-3', where Y is carboxy-dT (Glen Research cat#10-1035-90) and p is a 5'-terminal phosphate group with a photoprotection group (Glen Research cat# 10-4913-90) was obtained from DNA technology A/S Denmark. All oligonucleotides were produced using standard phosphoamidite chemistry. Oligo B was loaded with a functional entity comprising a cleavable linker as shown schematically below using the following protocols:

Synthesis of a Building block: N-Boc-Allyl-Glycine 0.12 mmol (26 mg) was dissolved in anhydrous acetonitrile and added 0.12 mmol (15 mg) 3-Amino-propionic acid methyl ester. The solution was cooled to 0° and added 1 mL of triethylamine and 0.12 mmol (45 mg) of HBTU. The reaction mixture was stirred over night at room temperature and evaporated to dryness. The remainder was dissolved in methanol (5 mL) and purified using reverse phase HPLC. (Yield: 53 % (0.063 mmol, 19mg)). The Boc protection group was removed by stirring the product in DCM containing 10 % TFA for 1 hour at room temperature. Allyl Glycine Beta Alanine Methylester was obtained as triflouroacetate in approx. 100 % yield (20 mg).

10 nmol of oligo B was dissolved in 50 µl of 200 mM Hepes-OH buffer pH 7.5 before addition of 20 µl of 100 mM of the chemical entity intermediate produced above in DMF, 20 µl 250 mM 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide (EDC) and 10 µl 100 mM of N-hydroxysuccinimide (NHS) in a total volume of 100 µl. The sample is incubated at 30 °C for 4 hours. Subsequently, oligo B comprising the loaded functional entity, i.e. the building block, is purified using HPLC.



The scheme above is a representation showing the synthesis of building block B comprising a DNA oligonucleotide sequence and a chemical entity.

The building blocks A and B were ligated to a template-primer complex having the sequence: 5'-pCTAGGTAGTAGACGATATCTCTACTACCTAGATGACATATCAAGT-3'. The synthesis is schematically shown on Fig. 7. 10 pmol of building block A and 10 pmol of building block B were pre-activated by removal of the 5'-photoprotection group using a Vilber-Lourmat trans-illuminator and subjecting the sample to UV-light at 312 nm for 30 seconds. Next the anticodon building blocks A and B were ligated to 10 pmol of template-primer complex using the TaKaRa ligation kit version 2.0. The three oligonucleotides were incubated in volume of 20 µl and were briefly denatured at 80 °C for 2 minutes before slowly cooling down to ambient temperature. An equivolume of TaKaRa ligation solution 1 was added and the sample incubated at ambient temperature for 1 hour. The ligation product was extracted twice in equivolumes of phenol before gelfiltration using Bio-rad microspin 6 columns. The ligation product was examined by electrospray-MS analysis (Bruker Inc.).

The ligation product was denatured by addition of 20 mM NaOH for 2 minutes at 25 °C in a total volume of 10 µl before addition of 50 µl of 9.5 M urea dissolved in 100 mM Hepes-OH buffer pH 7.5 (final conc). 5 µl of 500 mM EDC and 5 µl of 200 mM NHS were added and the sample incubated overnight at ambient temperature. The sample was purified using double gel-filtration before testing the cross-linking product on Electrospray-MS (Bruker Inc.). Successful crosslinking was observed by the removal of water from the ligation product (i.e. removal of 18 Dalton from the total molecular mass). Following cross-linking, the linker was cleaved by addition of 10 µl of a solution containing 25 mM iodine in THF and incubated at 30 °C for 1 hour. The sample was purified by double gel-filtration using bio-rad 6 spin-columns before testing the cross-linking product on electrospray-MS analysis (Bruker Inc.). Successful cross-linking and cleavage of the linker was observed as the addition of iodine to the total mass without elimination of the β-alanine building block (which have been cross-linked and transferred to the amino-nucleophile of building block A).

Claims

1. A method for synthesising a bifunctional complex comprising an encoded molecule and a template coding for one or more chemical entities which have participated in the synthesis of the encoded molecule, the method comprising the steps of

i) providing

a) a template comprising one or more codons

b) one or more building blocks having an anticodon associated with a chemical entity, and

c) a nucleic acid sequence associated with a reactive site,

ii) contacting the tempate with the one or more building blocks under conditions allowing for hybridisation between codons and anticodons,

iii) ligating at least one anticodon of a building block to the nucleic acid sequence associated with the reactive site, and

iv) reacting the chemical entity of the ligated building block with the reactive site under conditions where the ligation product is single stranded, to obtain a template-encoded reaction product.

2. The method according to claim 1, wherein the template comprises a flanking region and a priming region, said flanking region being complementary to the priming region allowing for a hairpin loop to be formed.

3. The method according to claim 1, wherein the anticodon is attached to a solid support.

4. The method according any of the preceding claims, wherein the nascent building block is attached to a solid support and worked up to the final building block remaining connected to said solid support.

5. The method according to any of the preceding claims, wherein the building blocks are incorporated stepwise.

6. The method according to any of the preceding claims, wherein two or more building blocks are hybridised to the template and subsequently ligated together to form a ligation product.

7. The method according to any of the preceding claims, wherein the anticodon is ligated to a nucleic acid using a ligase.

8. The method according to claim 1, wherein steps ii) through iv) is repeated as appropriate using a nascent complex as the template and anticodon(s) directed to a non-used codon in the building blocks to be incorporated.

9. The method according to any of the preceding claims, wherein a plurality of templates and building blocks are processed simultaneously or sequentially forming a library of complexes.

10. A library of different complexes, each complex comprising an encoded molecule and a template, which has encoded the chemical entities which has participated in the synthesis thereof, said library being obtainable by processing a plurality of different templates and a plurality of building blocks in accordance with any of the claims 1 to 9.

Abstract

Disclosed is a method for synthesising a bifunctional complex comprising an encoded molecule and a template coding for chemical entities which have participated in the synthesis of the encoded molecule. The method comprises the steps of providing: a template comprising one or more codons, one or more building blocks having an anti-codon associated with a chemical entity, and a nucleic acid sequence associated with a reactive site; contacting the tempate with the one or more building blocks under conditions allowing for hybridisation between codons and anticodons; ligating at least one anticodon of a building block to the nucleic acid sequence associated with the reactive site; and reacting the chemical entity of the ligated building block with the reactive site under conditions where the ligation product is single stranded, to obtain a template-encoded reaction product.

Fig. 1

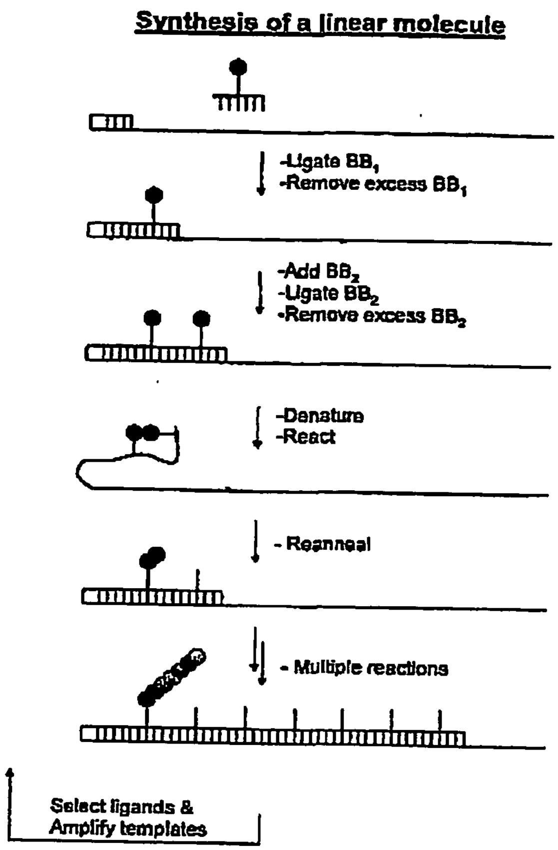

Synthesis of a linear molecule
-Ligate BB_1
-Remove excess BB_1
-Add BB_2
-Ligate BB_2
-Remove excess BB_2
-Denature
-React
- Reanneal
- Multiple reactions
Select ligands &
Amplify templates

Fig.2

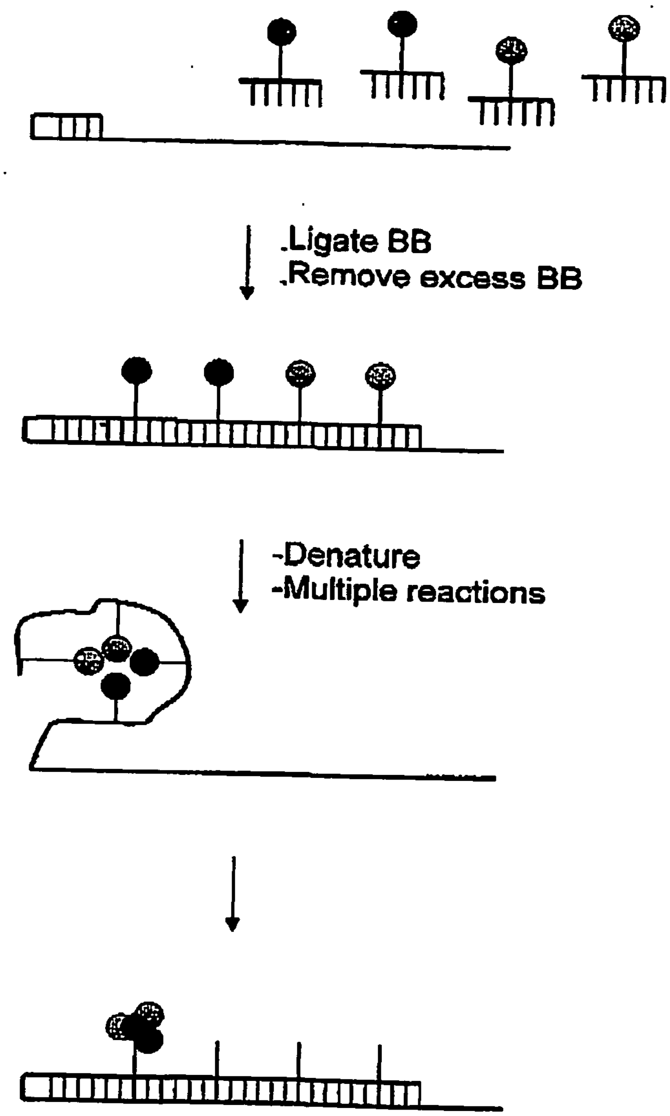


Fig. 3

```
   Codon 4  Codon 3  Codon 2  Codon 1
                                      GGAGCTGGTG T T
5' AGCAGTTTGTAGTAAGTACTTTGTAGTACCTCGACCAC  T T T
```

Anticodon-building block 7-mers

Anticodon-building block subtype 1: pTNNXNNA

Anticodon-building block subtype 2: pANNXNNT

X = Carboxy dT with attached building block

Modtaget

2 0 MRS. 2003

PVS

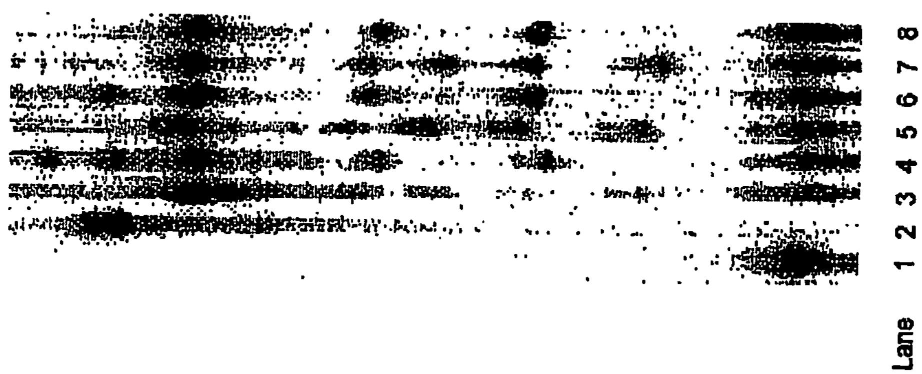
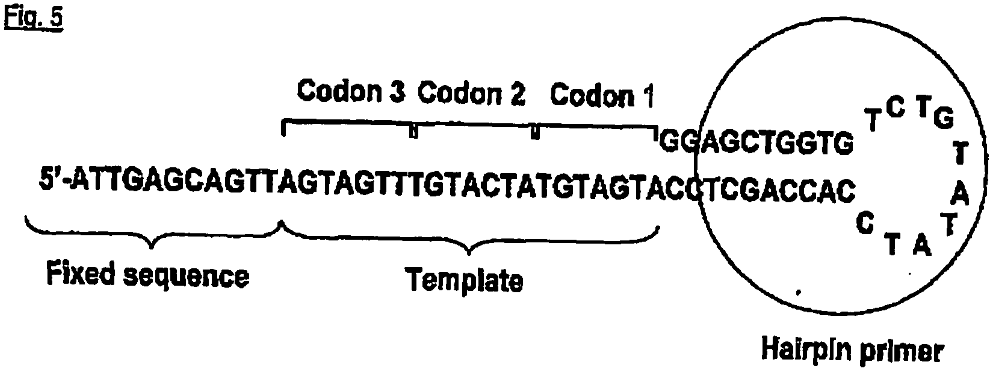


Fig. 4

Fig. 5



<u>Building block 7-mer oligonucleotides</u>

BB 1: pTACXACA complementary to codon 1

BB 2: pTACXACA complementary to codon 2

BB 3: pAACXACT complementary to codon 3

X = Carboxy dT

p = Phosphate

Fig. 6

```
Codon 4   Codon 3   Codon 2   Codon 1
                                      GGAGCTGGTG T T
5' AGCAGTTTGTAGTAAGTACTTTGTAGTACCTCGACCAC T T T
```

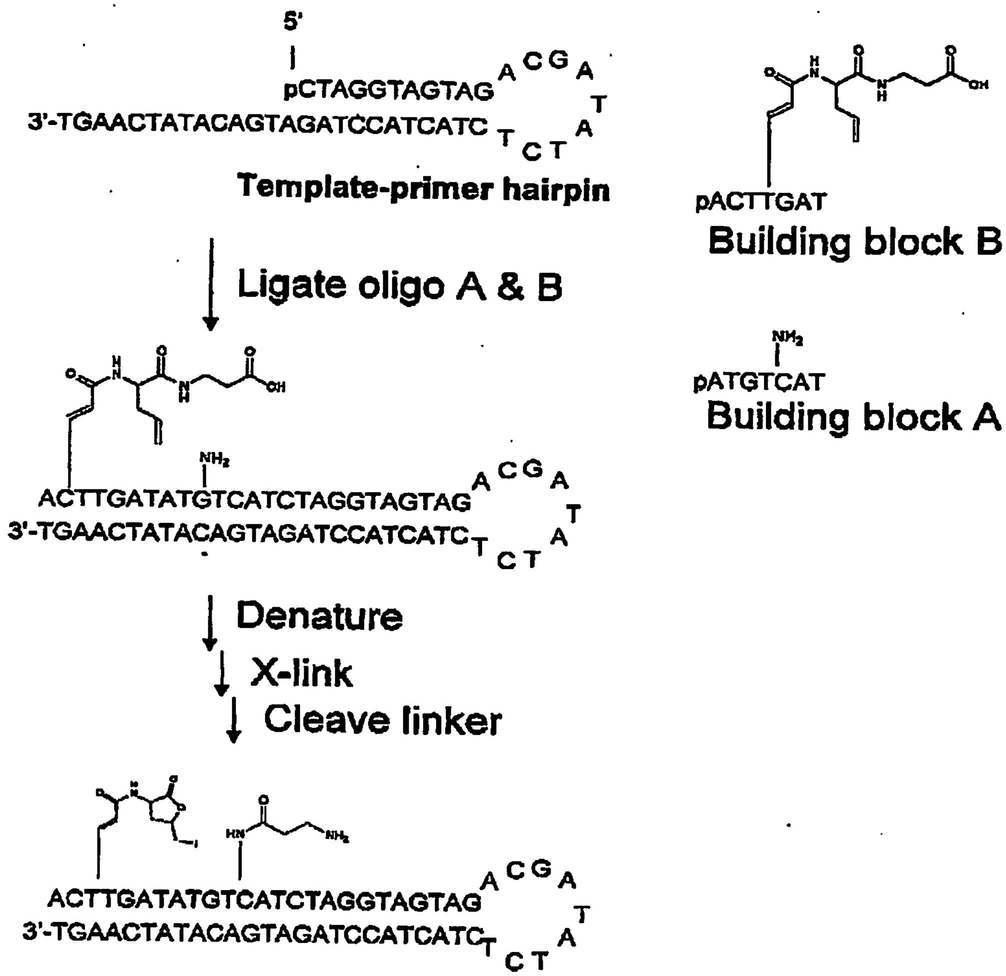
Anticodon-building block 7-mers

Anticodon-building block subtype 1: pTNNXNNA

Anticodon-building block subtype 2: pANNXNNT

X = Carboxy dT with attached building block

Fig.7


5'
pCTAGGTAGTAG
3'-TGAACTATACAGTAGATCCATCATC
Template-primer hairpin
pACTTGAT
Building block B
Ligate oligo A & B
pATGTCAT
Building block A
ACTTGATATGTCATCTAGGTAGTAG
3'-TGAACTATACAGTAGATCCATCATC
Denature
X-link
Cleave linker
ACTTGATATGTCATCTAGGTAGTAG
3'-TGAACTATACAGTAGATCCATCATC